

General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

A PROVISIONAL EFFECTIVE EVALUATION
WHEN ERRORS ARE PRESENT IN INDEPENDENT VARIABLES

L. S. Gurin



Translation of "Uslovno-effektivnoye Otsenivaniye pri Nalichii Oshibok v Nezavisimyykh Peremennyykh," Academy of Sciences USSR, Institute of Space Research, Moscow, Report Pr-753, 1982, pp. 1-22.

(NASA-TM-77211) A PROVISIONAL EFFECTIVE
EVALUATION WHEN ERRORS ARE PRESENT IN
INDEPENDENT VARIABLES (National Aeronautics
and Space Administration) 23 p
HC A02/MF A01

N83-33592

Unclass

CSCL 12A G3/64 28584

STANDARD TITLE PAGE

1. Report No. NASA TM-77211	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle A PROVISIONAL EFFECTIVE EVALUATION WHEN ERRORS ARE PRESENT IN INDEPENDENT VARIABLES		5. Report Date May 1983	
		6. Performing Organization Code	
7. Author(s) L. S. Gurin		8. Performing Organization Report No.	
		10. Work Unit No.	
9. Performing Organization Name and Address SCITRAN Box 5456 Santa Barbara, CA 93108		11. Contract or Grant No. NASA 3542	
		12. Type of Report and Period Covered Translation	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code	
13. Supplementary Notes Translation of "Uslovno-effektivnoye Otsenivaniye pri Nalichii Oshibok v Nezavisimyykh Peremennyykh", Academy of Sciences USSR, Institute of Space Research, Moscow, Report Pr-753, 1982, pp. 1 - 22.			
16. Abstract Algorithms are examined for evaluating the parameters of a regression model when there are errors in the independent variables. The algorithms are fast and the estimates they yield are stable with respect to the correlation of errors and measurements of both the dependent variable and the independent variables. ORIGINAL PAGE IS OF POOR QUALITY			
17. Key Words (Selected by Author(s))		18. Distribution Statement Unclassified - Unlimited	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 23	22. Price

A PROVISIONAL EFFECTIVE EVALUATION
WHEN ERRORS ARE PRESENT IN INDEPENDENT VARIABLES

L. S. Gurin
Space Research Institute, USSR Academy of Sciences

Submitted by Deputy Director V.M. Balebabov

Algorithms are examined for evaluating the parameters of a regression model when there are errors in the independent variables. The algorithms are fast and the estimates they yield are stable with respect to the correlation of errors and measurements of both the dependent variable and the independent variables.

1. Formulation of the Problem

We consider a model of a system (or physical phenomenon) of form:

/3*

$$y = f(x, \theta), \quad (1)$$

where:

$$x \in R^m$$

is the vector of the input quantities (independent variables or regressors); $\theta \in R^k$ is the vector of the parameters being evaluated; $y \in R^1$ is the output quantity (dependent variable).

The literature also discusses the case when there are errors in the model, i.e. in reality:

$$y = f(x, \theta) + \xi', \quad (2)$$

*Numbers in the margin indicate pagination in the foreign text.

ORIGINAL PAGE IS
OF POOR QUALITY

where ξ' is a random quantity that allows for the joint influence of many factors, having a certain law of distribution. But this case does not deal with an experiment and its errors. Hereafter we shall not discuss this case, intending (as well as the majority of authors) that the error of the model ξ' can be included in the measurement error of the quantity y .

As a result of the experiment, n sets of measured values are produced:

$$\bar{y}_i = y_i + \xi_i ; \quad \bar{x}_i = x_i + \eta_i ; \quad i = 1, 2, \dots, n \quad (3)$$

We must find an estimate of the parameter θ , i.e.

$$\hat{\theta}_n = \Phi_n(\bar{x}, \bar{y}) \quad (4)$$

The literature concerning the evaluation, numbering in the thousands, owes its plenitude to the fact that the formulation of the problem can be improved in various manners, depending on the assumptions made as to the type of function f , the class of algorithms Φ_n among which the optimal is chosen, the initial data as to the parameters and random factors, and the criteria by which the algorithms ϕ_n are compared.

The characteristic features of the provisional effective evaluation are the following:

a) only finite algorithms are considered as candidates for ϕ_n ;

b) the evaluations are compared in terms of not one, but several criteria;

c) one of the criteria of comparison of the estimates is the difficulty of realization (complexity) of the algorithm ϕ_n ;

d) the investigation favors the results of statistical modeling, as the asymptotic properties of finite algorithms and finite samples are not sufficient to make decisions as to the quality of the estimates.

The concept of a provisional effective evaluation was introduced in [1], where they considered the difficulty of the algorithms, along with precision criteria. A similar concept of a provisional optimal evaluation is examined in [2] in the solving of dynamic problems. Hereafter we shall add the stability (robustness) in the sense of [3,4] to the quality criteria of the estimates. The corresponding provisional effective estimates have been considered in [5]. In [6] both the robustness in the above sense and the robustness (guarantee) in the sense of [7] are considered.

/5

A general survey of the problem of provisional effective estimation as a multi-criterial problem of investigation of operations is given in [8]. In all the work on provisional effective estimation to the present it is assumed that the independent variables are measured without error.

As concerns the allowance for errors in the independent variables, more than a hundred articles have already been devoted to this, as well as sections in monographs (e.g. [9-16]). With very few exceptions (e.g. [17]), the errors in both the dependent and the independent variables in these works are assumed as non-correlated for the different measurements (although they may be correlated between the individual components of the input vector x [18]).

The aim of the present work is to obtain provisional effective estimates while allowing for correlated errors in the independent variables. Thus, we first require very simple estimates and, in the second place, we must allow for errors

(even correlated ones) in the independent variables. Thirdly, the estimates should be robust with regard to the correlation of errors, i.e. in the sense of [7]. Let us note at the outset, to avoid a later repetition, that the proposed estimates can also be made robust in the sense of [3,4] by including them in the two-level scheme of [5]. It is natural that, as in the general case of solving multi-criterial problems, the obtained estimates will not be optimal when examined for each individual criterion.

2. Examination of the Elementary Model

Let us examine an elementary linear model:

$$y = \theta x ; \quad m=1 ; \quad \kappa=1 \quad (5)$$

With regard to the errors ξ_i and η_i from (3) we make the following assumptions:

$$\left. \begin{aligned} E\xi_i = E\eta_i = 0 ; \quad \text{cov}(\xi_i, \eta_j) &= 0 ; \\ \text{cov}(\xi_i, \xi_j) &= \sigma_{\xi ij}^2 ; \quad \text{cov}(\eta_i, \eta_j) = \sigma_{\eta ij}^2 ; \\ i, j &= 1, 2, \dots, n \end{aligned} \right\} \quad (6)$$

We introduce a new quantity:

$$\bar{z} = \bar{y} - \theta \bar{x} \quad (7)$$

We then have:

$$E\bar{z}_i = 0 ; \quad \text{cov}(\bar{z}_i, \bar{z}_j) = \sigma_{\bar{z} ij}^2 = \sigma_{\eta ij}^2 + \theta^2 \sigma_{\xi ij}^2 \quad (8)$$

(following [17] we could also allow for a correlation of ξ and η ,

but we shall simplify the model even further for greater clarity in the later arguments).

According to [17], the estimate $\hat{\theta}$ is found by minimization of the quadratic form:

$$U = \bar{z}' \Sigma_z^{-1} \bar{z}, \quad (9)$$

where:

$$\bar{z}' = (\bar{z}_1, \dots, \bar{z}_n) ; \Sigma_z = \|\sigma_{z_{ij}}^2\|$$

We can use various methods of minimization, but in all cases a difficult algorithm results. In [17] the whole reduces to a solving of algebraic equations of high degree, while the coefficients of these equations, in turn, are found by computations with matrices of high order.

Since a simple algorithm is required for the provisional effective evaluation, we shall use the following approach. We denote the matrices:

$$\Sigma_y = \|\sigma_{y_{ij}}^2\| ; \Sigma_x = \|\sigma_{x_{ij}}^2\| \quad (10)$$

Then (8) gives:

$$\Sigma_z = \Sigma_y + \theta^2 \Sigma_x \quad (11)$$

We shall regard $\theta^2 \Sigma_x$ as small in comparison with Σ_y . We note that this condition may be ambiguous in meaning. It is fulfilled if the errors η are small in relation to those of ξ , which happens rather often (this explains the usually adopted assumption of their equaling zero). Along with this we may also consider the case when the errors ξ and η are comparable, but the parameter θ is small. Each case may be reduced to the other

by appropriate change in scale. Therefore, in future, we shall simply regard θ as small.

Under this assumption we have (I is the unit matrix):

$$\Sigma_z^{-1} = (I + \theta^2 \Sigma_y^{-1} \Sigma_x)^{-1} \Sigma_y^{-1} \approx \Sigma_y^{-1} - \theta^2 \Sigma_y^{-1} \Sigma_x \Sigma_y^{-1} \quad (12)$$

We introduce the designations:

$$\Sigma_y^{-1} = \|a_{ij}\| ; \Sigma_y^{-1} \Sigma_x \Sigma_y^{-1} = \|b_{ij}\| \quad (13)$$

Then (8) is rewritten as:

$$\left. \begin{aligned} U = \sum_{i,j} a_{ij} \bar{y}_i \bar{y}_j - \theta \left[\sum_{i,j} a_{ij} (\bar{x}_i \bar{y}_j + \bar{x}_j \bar{y}_i) \right] - \\ - \theta^2 \sum_{i,j} b_{ij} \bar{y}_i \bar{y}_j + \theta^2 \sum_{i,j} a_{ij} \bar{x}_i \bar{x}_j \end{aligned} \right\} \quad (14)$$

Here we have discarded terms containing θ to a degree higher than the second.

/8

Minimizing U and taking account of the symmetry of the matrices, we get:

$$\hat{\theta} = \frac{\sum_{i,j} a_{ij} \bar{x}_i \bar{y}_j}{\sum_{i,j} (a_{ij} \bar{x}_i \bar{x}_j - b_{ij} \bar{y}_i \bar{y}_j)} \quad (15)$$

Thus we have solved the first part of our problem: an elementary estimate is found that allows for correlated errors in the independent variables (as well as the dependent). But for this we have assumed the covariational matrices Σ_y and Σ_x as given. In order to have guaranteed estimates we shall use the following method (for the case of no errors in the indepen-

dent variables this method had been previously studied thoroughly and used by V. L. Gurin).

Let the covariational matrices Σ_y and Σ_x have the form:

$$\Sigma_y = \sigma_y^2 I + \Sigma_y(\lambda) ; \quad \Sigma_x = \Sigma_x(\lambda), \quad (16)$$

where $\Sigma_y(\lambda)$ and $\Sigma_x(\lambda)$ satisfy, by each of their elements, boundary conditions of type:

$$|\sigma_{yij}^2(\lambda)| \leq \lambda \sigma_{yij}^2(1) ; \quad \theta^2 |\sigma_{xij}^2(\lambda)| \leq \lambda \sigma_{xij}^2(1) \quad (17)$$

We have assumed that:

$$\|\sigma_{yij}^2(1)\| \quad \text{and} \quad \|\sigma_{xij}^2(1)\|$$

are given in advance.

An analysis of (15) shows that, under our assumptions, we can write:

$$\hat{\theta} = \theta - \theta \sum_i \frac{x_i}{\sum_i x_i^2} \eta_i + \sum_i \frac{x_i}{\sum_i x_i^2} \xi_i + O(\lambda), \quad (18)$$

where the term $O(\lambda)$ includes the terms that vanish when $\lambda = 0$, as well as those that contain η_i and ξ_i beyond the first power. We also note that the x_i values figuring in (18) are not the measured, but the actual values, while:

$$\theta x_i = y_i \quad .$$

Further, in identical manner we can write for the dispersion of the estimate:

$$\sigma^2(\hat{\theta}) = \sum_{i,j} \theta^2 \sigma_{xij}^2 \frac{x_i x_j}{(\sum_i x_i^2)^2} + \sum_{i,j} \sigma_{yij}^2 \frac{x_i x_j}{(\sum_i x_i^2)^2} + O(\lambda) \quad (19)$$

To obtain estimates that are robust in regard to correlation we must find the minimax value $\sigma^2(\hat{\theta}_\lambda)$, provided that the class of covariational matrices is determined by expression (17) with fixed value of λ , while the particular class of algorithms includes linear algorithms that are unbiased in the first approximation (i.e. when $\lambda \rightarrow 0$). In examining the corresponding game with nature we observe that, when a saddle point is present, we can replace the minimax with the maximin. On the other hand, we can consider in a first approximation that, for fixed covariational matrices, estimate (15) has the least dispersion in the class of these particular algorithms. We shall prove the existence of the saddle point by a construction, i.e. specifically indicating the appropriate strategies. We denote:

$$\left. \begin{aligned} \check{\sigma}_{xij}^2 &= \text{Sign}(x_i x_j) \frac{\lambda}{\theta^2} \sigma_{xij}^2(1); \\ \check{\sigma}_{yij}^2 &= \sigma_{yij}^2 + \text{Sign}(x_i x_j) \lambda \sigma_{yij}^2(1) \end{aligned} \right\} \quad (20)$$

and shall take this as the strategy of nature. Our strategy is obtained from expression (15) if a_{ij} and b_{ij} correspond to the covariational matrices (20). We denote the corresponding estimate by $\hat{\theta}$. The obtained pair of strategies will be the saddle point when, and only when, the familiar conditions are fulfilled:

$$\sigma^2(\hat{\theta}; \sigma_{xij}^2, \sigma_{yij}^2) \leq \sigma^2(\hat{\theta}; \check{\sigma}_{xij}^2, \check{\sigma}_{yij}^2) \leq \sigma^2(\hat{\theta}_\lambda; \check{\sigma}_{xij}^2, \check{\sigma}_{yij}^2) \quad (21)$$

The inequality on the right of (21) follows from our previous discussion. In regard to the left inequality, this is obtained as follows. As is evident from formula (19), when $\lambda = 0$, the sign of the partial derivative of $\sigma^2(\hat{\theta})$ with respect to σ_{xij}^2 or σ_{yij}^2 agrees with the sign of $x_i x_j$. Consequently, a certain neighborhood of the point $\lambda = 0$ exists where this sign conformity is maintained. In view of the approximations in our theory,

i.e. the assumption of small values of λ , we obtain the requisite inequality from this fact and the definition (20).

Let us formulate the definitive results. If we know, when solving problem (5)-(6), that the covariational matrices of the measurement errors Σ_y and Σ_x satisfy the conditions (16)-(17) and λ is sufficiently small, we can use the estimate (15) as the provisional effective estimate $\hat{\theta}$, where a_{ij} and b_{ij} have been found from (13) for matrices (20).

In the spirit of the provisional effective approach to the evaluation, this result requires a confirmation by statistical modeling.

/11

We make some further remarks.

1. The value θ figuring in (20) is previously unknown, although it can be specified on the basis of a solution of analogous problems by methods that do not allow for errors in the independent variables. The precision with which θ is given is not especially important in light of the fact that λ is also provisionally specified. We need only realize that the result will be more accurate as λ is smaller. In regard to the values of x_i, x_j also figuring in (20), these can be replaced by \bar{x}_i, \bar{x}_j , as it is unlikely that this affects the sign of the product.

2. In forming the elements of the matrices $\check{G}_{xij}^z, \check{G}_{yij}^z$ we must check to be sure that the obtained matrices are covariational; for Σ_y this will be automatically fulfilled when λ is small, since the matrix which differs little from a diagonal matrix with positive elements will be positively determined. For Σ_x we can use a similar argument in constructing the boundary conditions.

ORIGINAL PAGE IS
OF POOR QUALITY

3. In the boundary conditions on the set of strategies of nature (16)-(17) it is assumed that the quantities:

$$\sigma_{xij}^1(i), \sigma_{yij}^1(i)$$

as well as:

$$\sigma_{xij}^1(\lambda), \sigma_{yij}^1(\lambda)$$

are not connected by any sort of relation. If such correlations obtain, the form of the estimates will be slightly modified. Let us consider several examples.

a) Usually the errors are in the form of a stationary random sequence, i.e. there obtains:

$$\left. \begin{aligned} \sigma_{xij}^2(i) &= \sigma_{xik}^1(i); \quad \sigma_{yij}^2(i) = \sigma_{yik}^1(i); \quad \sigma_{xij}^2(\lambda) = \sigma_{xik}^1(\lambda); \\ \sigma_{yij}^2(\lambda) &= \sigma_{yik}^1(\lambda); \quad k = |i-j| \end{aligned} \right\} \quad (22)$$

In this case the elements of the matrices

$$\check{\sigma}_{xij}^1 \quad \text{and} \quad \check{\sigma}_{yij}^1$$

also depend solely on k and the largest or smallest possible ones are chosen, depending on the sign:

/12

$$\psi(k) = \sum_{i=1}^{n-k} x_i x_{i+k} \quad (23)$$

b) The set of matrices can be narrowed down not only by multiplying with a certain factor λ , but also in other ways. For example, let the matrices be stationary (as in §a) and:

$$\sigma_{xk}^2(i) = \sigma_x^1(i) \rho_x^k(i); \quad \sigma_{yk}^2(i) = \sigma_y^1(i) \rho_y^k(i); \quad (k=1, 2, \dots, n-1) \quad (24)$$

i.e. the boundary conditions form Markov matrices. It is then more convenient to introduce the factor λ that reduces the set of matrices in following fashion:

$$\theta^2 |G_{xk}^1(\lambda)| \leq \lambda G_x^1(1) \rho_x^{\frac{k}{\lambda}}(1); |G_{yk}^1(\lambda)| \leq G_y^1 \rho_y^{\frac{k}{\lambda}}(1); 0 < \lambda \leq 1 \quad (25)$$

3. Example

Let us have the relation:

$$\tilde{y} = \theta_0 + \theta \tilde{x} \quad (26)$$

and it is necessary to estimate θ_0 and θ . In order to convert the problem to model (5) we shall consider that \tilde{x} is measured in an even number $n = 2\ell$ of points at identical distances Δ (with precision down to the errors in assignment of \tilde{x}). We denote:

$$y = \tilde{y} - y_{cp}; \quad x = \tilde{x} - x_{cp}, \quad (27)$$

where y_{cp} and x_{cp} are the arithmetic means of the measured values of \tilde{y} and \tilde{x} . Then x and y are related by (5).

We shall further employ the above-explained results for the estimation of θ . We first make two remarks.

13

1. In connection with the allowed correlation of measurement errors and the presence of such errors in the measurement of \tilde{x} , such a conversion of model (26) to model (5) is less accurate than the simultaneous evaluation of θ_0 and θ , which can be done in our conditions by using the algorithm explained in the next section. But here we shall disregard it and thus obtain an estimate for θ_0 (after estimating θ) in the form:

$$\hat{\theta}_0 = y_{cp} - \hat{\theta} x_{cp} \quad (28)$$

2. Whereas at the outset there was given a class of allowable covariational matrices for \tilde{y} and \tilde{x} , this class will differ for y and x and the conversion can be easily done. We shall not concern ourselves with this here. We merely note that, even for uncorrelated measurement errors of x and y , the measurement errors of x and y will be correlated [sic!].

Thus, let us assume that we are already at the place discussed in the preceding section. To be specific, we assume that the boundary conditions on the covariational matrices have the form (25) with fixed value of λ . What is the physical meaning of this? We are assuming an arbitrary dispersion of errors y with a sufficiently small correlation, although the smaller the correlation the more accurate the analysis. The correlation matrix that majorizes the modulus is a Markov matrix, although in reality the errors may represent e.g. a stationary process of autoregression of any given order. Since λ has been chosen for satisfactory approximation in the theory, we denote for brevity:

$$\rho_y^{\lambda}(1) = \rho_y \quad (29)$$

The same can be said of x and we denote:

$$\rho_x^{\lambda}(1) = \rho_x \quad (30)$$

Here, however, we must add that the dispersion of errors y cannot be arbitrary, as follows from the difference between formulas (25) for y and x . For simplicity we shall consider hereafter that, as stipulated above, the scale of x is chosen so that (25) is satisfied by the agency of θ^2 , whereas $\sigma_x^2(1)$ and even $\sigma_x^2(\lambda)$ may be comparable to σ_y^2 .

Thus, to obtain guaranteed estimates we everywhere replace σ_x^2 and σ_y^2 by their upper limits. In regard to the nondiagonal

elements of the covariational matrices:

$$\check{\Sigma}_x \text{ and } \check{\Sigma}_y,$$

these are taken as the largest or smallest values, depending on the sign of $\psi(k)$ from formula (23). We note that, as already mentioned above in regard to (20), instead of x_i we can insert in (23) their real values or any others sufficiently close to them. From the remarks at the outset of this section it follows that we can take:

$$x_i = \Delta(2i - 2\ell -); \quad i = 1, 2, \dots, 2\ell \quad (31)$$

Hence we obtain:

$$\psi(k) = \sum_{i=1}^{2\ell-k} x_i x_{i+k} = 4(2\ell-k) \left[2\ell^2 - 2\ell k - k^2 - \frac{1}{2} \right], \quad (32)$$

i.e. $\psi(k) > 0$ when:

$$k < -\ell + \sqrt{3\ell^2 - \frac{1}{2}} = c(\ell) \quad (33)$$

and $\psi(k) < 0$ when $k > c(\ell)$. We can assume with sufficient accuracy that:

$$c(\ell) \approx \ell(\sqrt{3} - 1) \approx 0,73\ell \quad (34)$$

The obtained matrices:

$$\check{\Sigma}_x \text{ and } \check{\Sigma}_y$$

will have negative elements only far away from the main diagonal. Since it follows from (25) that these elements will be extremely small, the positive determination of the matrices is assured. However it is not exactly easy to invert them by computer. We shall therefore take the following step. Considering the small

nature of these elements, we shall replace their sign by a positive, which little affects the later results. But now the matrices become Markov and they can be inverted analytically. To obtain a_{ij} and b_{ij} from (13) and estimate (15), we have only to calculate the matrix product:

$$\check{\Sigma}_y^{-1} \check{\Sigma}_x \check{\Sigma}_y^{-1},$$

the elements of which have an analytic expression. This can also be done analytically (we shall not give here the corresponding formulas), but even when the computer is used it is much less time-consuming than a numerical matrix inversion.

4. Investigation of a More Complex Model

Let us have, instead of (5), a model of type:

$$z = \theta' x, \quad (35)$$

where:

$$\theta' = (-\theta_1, \dots, -\theta_K, \theta_{K+1}) ; \quad \theta_{K+1} = 1 ; \quad (36)$$

$$x' = (x_1, \dots, x_K, x_{K+1}) ; \quad x_{K+1} = y \quad (37)$$

For the given covariant error matrices this version has also been considered in [17] in nonsimplified form. In this case the quadratic form (9) is much more complicated:

$$\sigma_{zij}^2 = \theta' \Sigma_{ij} \theta, \quad (38)$$

where:

$$\Sigma_{ij} = \|\sigma_{ij}^2(\mu, \nu)\|$$

and

$$\sigma_{ij}^2(\mu, \nu) = \text{COV}(x_{\mu i}, x_{\nu j}) \quad (39)$$

In (39) i, j are the numbers of the measurements, while μ, ν are the numbers of the variables, i.e. $i, j = 1, 2, \dots, n$; $\mu, \nu = 1, 2, \dots, k+1$.

To simplify the problem, by analogy with the case $k = 1$, explained above, we shall take $\theta_1, \dots, \theta_k$ as small (which, as in the mentioned case, can be done by choice of the scales when there are small errors in x_1, \dots, x_k). To simplify the later computations we assume that the errors of the various variables are not correlated with each other (this assumption need not be used: it merely complicates the formulas but introduces no essential modifications). Then:

/16

$$\sigma_{ij}^2(\mu, \nu) = 0 \text{ when } \mu \neq \nu$$

and (38) becomes:

$$\sigma_{ij}^2 = \sum_{\nu=1}^k \theta_{\nu}^2 \sigma_{ij}^2(\nu) + \sigma_{ij}^2(k+1), \quad (40)$$

i.e. instead of (11) we have:

$$\Sigma_z = \Sigma_{k+1} + \sum_{\nu=1}^k \theta_{\nu}^2 \Sigma_{\nu} \quad (41)$$

Instead of (12) we get:

$$\Sigma_z^{-1} \approx \Sigma_{k+1}^{-1} - \Sigma_{k+1}^{-1} \left(\sum_{\nu=1}^k \theta_{\nu}^2 \Sigma_{\nu} \right) \Sigma_{k+1}^{-1} \quad (42)$$

or:

$$\Sigma_z^{-1} \approx \Sigma_{k+1}^{-1} - \sum_{\nu=1}^k \theta_{\nu}^2 \left[\Sigma_{k+1}^{-1} \Sigma_{\nu} \Sigma_{k+1}^{-1} \right] \quad (43)$$

We then construct the quadratic form U , similar to (14), and discard terms with θ_{ν} above the second power. As a result we get:

/17

$$\left. \begin{aligned}
 U &= \sum_{i,j} (\bar{x}_{k+1,i} - \sum_{v=1}^k \theta_v \bar{x}_{v,i}) (\bar{x}_{k+1,j} - \sum_{v=1}^k \theta_v \bar{x}_{v,j}) (a_{ij} - \sum_{v=1}^k b_{vij} \theta_v^2) \approx \\
 &\approx \sum_{i,j} a_{ij} \bar{x}_{k+1,i} \bar{x}_{k+1,j} - \sum_{i,j} \sum_{v=1}^k \theta_v a_{ij} (\bar{x}_{v,i} \bar{x}_{k+1,j} + \bar{x}_{v,j} \bar{x}_{k+1,i}) + \\
 &+ \sum_{i,j} \sum_{\mu, v=1}^k a_{ij} \bar{x}_{v,i} \bar{x}_{\mu,j} \theta_v \theta_\mu - \sum_{i,j} \sum_{v=1}^k b_{vij} \bar{x}_{k+1,i} \bar{x}_{k+1,j} \theta_v^2
 \end{aligned} \right\} \quad (44)$$

Minimizing U we obtain (in view of the symmetry of the matrices

$$[a_{ij}] \text{ and } [b_{vij}])$$

the set of equations:

$$B\theta = c, \quad (45)$$

where the matrix B is determined as follows in terms of its elements ($\delta_{ij} = 0$ when $i \neq j$, $\delta_{ij} = 1$):

$$b_{v\mu} = \bar{x}'_v \Sigma_{k+1}^{-1} \bar{x}_\mu - \delta_{v\mu} \bar{x}'_{k+1} [\Sigma_{k+1}^{-1} \Sigma_v \Sigma_{k+1}^{-1}], \quad (46)$$

while the vector c has components:

$$c_v = \bar{x}'_v \Sigma_{k+1}^{-1} \bar{x}_{k+1} \quad (47)$$

From this we obtain the estimate:

$$\hat{\theta} = B^{-1}c \quad (48)$$

Estimate (48), just as (15) (obtained from (48) in the particular case $k = 1$), is simplified; the matrix B, which must be inverted, has order k, i.e. it presents no computational difficulties. Moreover, in the entire calculation process it is only necessary to invert the high-order matrix Σ_{k+1} once.

**ORIGINAL PAGE IS
OF POOR QUALITY**

But we have already seen in the case of $k = 1$ that this usually can also be done analytically (especially when speaking of guaranteed estimates).

/18

Moving on to construct provisionally effective guaranteed estimates we note that in this case there will be one significant difference from the case $k = 1$. The penalty function in the game with nature should be scalar, while the precision of the estimate $\hat{\theta}$ is expressed by its covariational matrix. Therefore we shall examine not the estimate $\hat{\theta}$, but the estimate of a certain linear form $\beta'\hat{\theta}$, e.g. the predicted value of x_{k+1} for given values of $x' = (x_1, \dots, x_k)$. Without repeating in detail the remarks made in the case $k = 1$, we merely give the analogs of the respective formulas.

Denoting by \bar{x} the measurement result matrix of the independent variables, instead of formula (18) that obtained in the case $k = 1$ we obtain the following expression for the form $\beta'\hat{\theta}$:

$$\begin{aligned} \beta'\hat{\theta} = & \beta'\theta + \beta'(x'x)^{-1}x'\delta_{k+1} + \beta'[(x'x)^{-1}\delta x'x_{k+1} - \\ & - 2(x'x)^{-1}x'\delta x(x'x)^{-1}x'x_{k+1}] + \dots \end{aligned} \quad (49)$$

Here:

$$\delta_{k+1} = \bar{x}_{k+1} - x_{k+1}$$

(i.e. this corresponds to the vector ξ in (18)), while $\delta x = \bar{x} - x$ is the measurement error matrix for the independent variables; the terms not written out can be ignored when λ and θ have correspondingly small values.

The choice of the largest or smallest of the values of $\sigma_{ij}^2(v)$ as candidate for $\hat{\sigma}_{ij}^2(v)$ will now depend on the sign of

the product of the coefficients of δ_{vi} and δ_{vj} in expression (49). In order to find the coefficient of δ_{vi} we may proceed as follows. By C_{ij} we denote the matrix that has a solitary nonzero element equal to unity at the intersection of the i -th row and the j -th column. Then the sought coefficient a_{vi} is equal to:

$$a_{vi} = \beta' [(x'x)^{-1} C_{vi} x_{k+1} - 2(x'x)^{-1} x' C_{vi} (x'x)^{-1} x' x_{k+1}] \quad (50)$$

This refers to the case of $v = 1, 2, \dots, k$. For $v = k+1$, of course:

/19

$$a_{k+1,i} = \beta' (x'x)^{-1} x' c_i, \quad (51)$$

where the vector column c_i has unity in the i -th position, while the other elements are zeroes. For each specific problem the choice of $\hat{S}_{ij}^{(v)}$ can be done by computer. The only operation of matrix inversion will be that of the matrix $x'x$, the order of which is k , i.e. small.

All the remarks made for the case of $k = 1$ also apply here.

5. Confirmation of the Hypotheses

The findings can also be used to check the hypotheses. Thus, in the case of $k = 1$, we can obtain a provisionally effective robust criterion to check the hypothesis of the absence of a trend for the quantity y as x varies, by using expression (15). In order to investigate the statistical characteristics of this criterion we must assign boundary conditions on the covariational error matrices for the measurements of x and y , find the coefficients a_{ij} and b_{ij} , and then obtain for $\theta = 0$ the distribution function (or its quantiles) of the criterion $\hat{\theta}$ as determined by formula (15). This is a rather elaborate task, but it is only solved once and the results of its solution can be summarized in tables.

In order to illustrate by the simplest example the difference between the proposed criterion and the existing (which does not allow for errors in the measurement of the independent variable), let us assume that the errors are not correlated and:

$$\sigma_x^2 \leq k \sigma_y^2 \quad .$$

When θ is small, the entire theory is accurate (and k need not be small in this case!) and we adopt the guaranteed value:

$$\sigma_x^2 = k \sigma_y^2 \quad .$$

Then the criterion $\hat{\theta}$ becomes:

$$\hat{\theta} = \frac{\sum_i \bar{x}_i \bar{y}_i}{\sum_i (\bar{x}_i^2 - k \bar{y}_i^2)} \quad (52)$$

(we recall that the measurements of \bar{x}_i and \bar{y}_i are not centered). As k increases, so does $\hat{\theta}$. Thus, if x is indeed measured with errors, the use of a criterion that does not take this into account increases the likelihood of an error of the second kind (we assume that $\theta = 0$ is the null hypothesis).

References

1. Gurin, L. S., Tsoy, K. A., "Applications of Statistical Modeling in the Elaboration and Comparative Analysis of Methods of Parameter Estimation" in *Obrabotka kosmicheskoy informatsii* [Processing of Space Data], Nauka, Moscow, 1976, p. 36-57.
2. Pugachev, V. S., "A Generalization of the Theory of Provisionally-Optimal Estimation and Extrapolation", *Doklady AN SSSR*, 262/3, 1982, p. 535-538.
3. Huber, P. J., "Robust Estimation of a Location Parameter", *Ann. Math. Statist.*, 35/1, 1964, p. 73-101.
4. Yershov, A. A., "Stable Methods of Parameter Evaluation (Survey)", *Avtomatika i telemekhanika*, 8, 1978, p. 66-100.

5. Gurin, L. S., Tsoy, K. A., "Stable (Robust) Provisionally-Effective Estimates" in Obrabotka informatsii, poluchayemoy po programme "Interkosmos" [Processing Data Acquired in the Interkosmos Program], Nauka, Moscow, 1982, p. 229-242.

6. Gurin, V. I., "The Characteristics of a Sample Median and Twofold-Robust Estimation" in Statisticheskiye izmereniya i primeneniye mikromashinnykh sredstv v izmereniyakh [Statistical Measurements and Applications of Microcomputers in Mensuration], Abstracts of Reports of the All-Union Symposium at Vil'nyus, 1982, p. 68-72.

7. Bakhshiyani, B. Ts., Nazirov, R. R., El'yasberg, P. Ye., Opredeleniye i korrektsiya dvizheniya [The Determination and Correction of Movement], Nauka, Moscow, 1980, 360 p.

8. Gurin, L. S., "Estimation of Parameters as a Problem in Operation Theory", in Metody obrabotki izmereniy [Measurement Processing Methods], Supplement, Mudrov, V. I., Kushko, V. L., Radio i svyaz', Moscow, in press.

9. Kendal, M. Dzh., Styuart, A., Statisticheskiye vyvody i svyazi [Statistical Deductions and Correlations], Nauka, Moscow, 1973, 900 p.

10. Seber, Dzh., Lineynyy regressionnyy analiz [Linear Regression Analysis], Mir, Moscow, 1980, 456 p.

11. Fedorov, V. V., Teoriya optimal'nogo eksperimenta [The Theory of an Optimal Experiment], Nauka, Moscow, 1971, 312 p.

12. Demidenko, Ye. Z., Lineynaya i nelineynaya regressii [Linear and Nonlinear Regressions], Finansy i statistika, Moscow, 1981, 302 p.

13. Malenkov, E., Statisticheskiye metody ekonometrii [Statistical Methods of Econometry], Statistika, Moscow, 1, 1975, 424 p; 2, 1976, 324 p.

14. Klepikov, N. P., Sokolov, S. N., Analiz' i planirovaniye eksperimentov metodom maksimuma pravdopodobabiya [Analysis and Planning of Experiments by the Maximum Similitude Method], Nauka, Moscow, 1964, 184 p.

15. Nalimov, V. V., Chernova, I. A., Statisticheskiye metody planirovaniya ekstremal'nykh eksperimentov [Statistical Methods of Planning Extreme Experiments], Nauka, Moscow, 1965, 340 p.

16. Zel'ner, A., Bayesovskiy metody v ekonometrii [Bayesian Methods in Econometry], Statistika, Moscow, 1980, 438 p.

17. Sprent, P., "Generalized Least-Squares Approach to Linear Functional Relationships", J. Roy. Statist. Soc., B., 28, 1966, p. 278-297.

18. Fedorov, V. V., "Regression Analysis When There are Errors in the Determination of the Predictor" in Voprosy kibernetiki, ed. by V. V. Nalimov, 47, Mathematical-Statistical Methods of Analysis and Experimental Planning, Nauka, Moscow, 1978, p. 69-75.